UNIVERSITÉ DE MONTRÉAL

CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, SUCCURSALE A, MONTREAL, QUEBEC H3C 3J7

L'ALGORITHME DE KARMARKAR

par

Jacques Desrosiers

Cette étude a été publiée grâce à une subvention du fonds F.C.A.R. pour l'aide et le soutien à la recherche.

Université de Montréal
Centre de recherche sur les transports - Publication #403
Mars 1986

87-05578/

Dépôt légal - 3e trimestre 1986

TABLE DES MATIÈRES

1. Formulation du problème et hypothèses de départ. 1

2. La transformation projective T et la fonction de potentiel f . . . 2

3. Optimisation du problème transformé. 3

4. Quelques démonstrations. 7

   4.1 Les rayons des sphères circonscrite et inscrite 7

   4.2 Approximation du premier ordre de g(y). 8

   4.3 Théorème 5. 8

   4.4 Théorème 3. 9

   4.5 Théorème 4. 10

   4.6 Théorème 2. 13

   4.7 Théorème 1. 13

5. Extensions au modèle 14

   5.1 Discussion des hypothèses de départ 14

   5.2 Méthode proposée. 16

   5.3 Point admissible initial. 20

6. Complexité 21

7. Bibliographie. 21

Figure 1 : Rayons des spères circonscrites et inscrites. 22

SOMMAIRE

Ce texte présente les principaux résultats théoriques de l'algorithme de Karmarkar pour la programmation linéaire. La complexité de cet algorithme est polynomiale et requiert $O(n^{3.5} L \ln L \ln \ln L)$ opérations en comparaison de $O(n^4 L \ln L \ln \ln L)$ pour l'algorithme ellipsoïdal de Khachiyan, où n est le nombre de variables et L est le nombre de bits en input. L'algorithme consiste en l'application répétée d'une transformation projective suivie de l'optimisation sur une sphère. Dans un but pédagogique, le détail de plusieurs démonstrations est donné.

Mots clés : programmation linéaire, transformation projective.

1. FORMULATION DU PROBLÈME ET HYPOTHÈSES DE DÉPART

Le problème à résoudre par l'algorithme de Karmarkar [1] est le suivant :

$$\begin{aligned} &\min c^T x \\ &\text{sujet à } Ax = 0 \\ &\qquad e^T x = 1 \\ &\qquad x > 0 \end{aligned}$$

où $c, x, \in \mathbb{R}^n$ et A de dimension m x n.

Les hypothèses sur le modèle sont :

H1 : Le domaine réalisable est l'intersection d'un espace affin $\Omega = \{x \mid Ax=0\}$ et d'un simplexe $S = \{x \mid e^T x=1,\ x \geq 0\}$.

H2 : Le système d'équations définissant Ω est homogène, i.e. le membre droit est nul.

H3 : Le minimum de la fonction objectif $c^T x$ est zéro.

H4 : Le problème est réalisable et $x^0 = e/n = (1/n,\ldots,1/n)^T$ est le point admissible initial.

H5 : Un paramètre de terminaison q est donné et on cherche un x admissible $(x \in \Omega \cap S)$ tel que $\frac{c^T x}{c^T x^0} < 2^{-q}$.

Si le paramètre $q = O(L)$ où L est fonction de la longueur de l'input, la solution approchée peut être convertie en une solution exacte, comme dans la méthode de l'ellipsoïde.

La discussion des hypothèses et l'extension à un modèle plus général sont présentées à la section 5.

2. LA TRANSFORMATION PROJECTIVE T ET LA FONCTION DU POTENTIEL f

Soit $a > 0$ un point admissible pour $\Omega \cap S = \{x \mid Ax=0, e^T x=1, x \geq 0\}$ et $D_a = \text{diag}\,[a_1, a_2, \ldots, a_n]$ une matrice diagonale. On définit la transformation projective $T_a : \mathbb{R}^n \to \mathbb{R}^n$ par

$$T_a(x) = y = \frac{D_a^{-1}x}{e^T D_a^{-1}x}, \qquad \text{i.e.} \quad y_j = \frac{x_j/a_j}{\sum_{j=1}^{n}(x_j/a_j)} \qquad j=1,\ldots,n.$$

Cette transformation possède les propriétés suivantes :

1. $e^T y = \dfrac{e^T D_a^{-1} x}{e^T D_a^{-1} x} = 1$.

2. si $x \in \{x \mid e^T x=1, x \geq 0\}$, alors $T_a(x) = y \in \{y \mid e^T y=1, y \geq 0\}$ car $a>0$ et $e^T D_a^{-1} x = \sum_{j=1}^{n}(x_j/a_j)$ est un scalaire strictement positif.

3. T_a applique le point a sur le centre du simplexe dans l'espace transformé : $T_a(a) = e/n$.

4. $$e^T D_a y = \frac{e^T D_a D_a^{-1} x}{e^T D_a^{-1} x} = \frac{e^T x}{e^T D_a^{-1} x} = \frac{1}{e^T D_a^{-1} x}$$

   d'où la transformation inverse T_a^{-1} est donnée par $T_a^{-1}(y) = x = \dfrac{D_a y}{e^T D_a y}$.

   La transformation projective T_a est une bijection.

5. $$Ax = 0 \Leftrightarrow \frac{AD_a y}{e^T D_a y} = 0 \Leftrightarrow AD_a y = 0 .$$

   Le système d'équations dans l'espace transformé devient :

$$AD_a y = 0$$
$$e^T y = 1$$
$$y > 0.$$

À la suite de cette transformation projective, la fonction objectif $c^T x$ devient $\dfrac{c^T D_a y}{e^T D_a y}$ qui n'est plus linéaire. On remarquera toutefois que

$$c^T x = 0 \Leftrightarrow \frac{c^T D_a y}{e^T D_a y} = 0 \Leftrightarrow c^T D_a y = 0.$$

On introduit ici la fonction de potentiel f donnée par $f(x) = \sum_{j=1}^{n} \ell n \left(\dfrac{c^T x}{x_j}\right)$. Applicant la transformation projective, on obtient

$$\begin{aligned} f(x) &= \sum_{j=1}^{n} \ell n \left(\frac{c^T x}{x_j}\right) = \sum_{j=1}^{n} \ell n \left(\frac{c^T D_a y / c^T D_a y}{a_j y_j / c^T D_a y}\right) \\ &= \sum_{j=1}^{n} \ell n \left(\frac{c^T D_a y}{y_j}\right) - \sum_{j=1}^{n} \ell n\, a_j \\ &= g(y). \end{aligned}$$

À une constante près, la fonction f(x) est transformée en une fonction g(y) du même type : la fonction de potentiel est invariante sous une transformation projective. On verra au théorème 1 que toute diminution de $c^T x$ peut être obtenue par une diminution adéquate de la fonction de potentiel f(x).

3. OPTIMISATION DU PROBLÈME TRANSFORMÉ

À chaque itération de l'algorithme, on possède un point $x^k > 0$ admissible pour $\Omega \cap S$; on cherche un nouveau point $x^{k+1} > 0$ admissible faisant décroître la fonction de potentiel f(x). Dans la suite de l'exposé, on notera par T_k la transformation projective à partir du point x^k et par $D_k = \text{diag}\,[x_1^k, x_2^k, \ldots, x_n^k]$ la matrice diagonale associée. La procédure est illustrée de la façon suivante :

$$x^k \xrightarrow{T_k(x^k)} e/n; \quad \begin{array}{l}\text{optimisation}\\ \text{approchée en y}\\ \text{du problème}\\ \text{transformé}\end{array} \quad y^k \xrightarrow{T_k^{-1}(^k y)} x^{k+1}.$$

L'optimisation approchée en y du problème transformé

$$\min g(y) = \sum_{j=1}^{n} \ell n \left(\frac{c^T D_k y}{y_j}\right) - \sum_{j=1}^{n} \ell n\, x_j^k$$

$$\text{sujet à } AD_k y = 0$$

$$e^T y = 1$$

$$y > 0$$

se fait sous deux aspects. D'une part, les contraintes de non-négativité $y > 0$ sont remplacées par la plus grande sphère centrée en e/n inscrite dans le simplexe défini par $\{y \mid e^T y = 1,\ y > 0\}$. Le rayon de cette sphère est $r = \frac{1}{\sqrt{n(n-1)}}$ (voir section 4.1) et on optimise à l'intérieur seulement, i.e. dans la sphère $B(e/n, \alpha r)$ avec $0 < \alpha < 1$, de telle sorte que le point obtenu y^k aura toutes ses composantes strictement positives. D'autre part, on utilise l'approximation du premier ordre de la fonction g(y) : on obtiendra ainsi l'objectif linéaire $c^T D_k y$ (voir section 4.2). Avec ces deux approximations, la solution approchée y^k est alors obtenue analytiquement (voir section 4.3).

A partir du point e/n, on cherche un point $y = e/n + h$ dans l'espace transformé, solution de

$$\min_h g(e/n + h) \quad \Longleftrightarrow \quad \min_h g(e/n + h)$$

$$\text{sujet à } AD_k(e/n + h) = 0 \qquad \text{sujet à } AD_k h = 0 \quad (AD_k e/n = 0)$$

$$e^T(e/n + h) = 1 \qquad e^T h = 0 \quad (e^T e/n = 1)$$

$$(e/n + h) > 0 \qquad (e/n + h) > 0$$

En utilisant les deux aspects de l'approximation,

$$(e/n + h) \in B(e/n, \alpha r),\ 0 < \alpha < 1 \equiv h^T h < \alpha^2 r^2,\ 1 < \alpha < 1$$

$$\min_h g(e/n + h) \simeq \min_h g(e/n) + \nabla_y g(e/n) \cdot h \equiv \min_h c^T D_k h$$

et en notant la matrice des contraintes par $B = \begin{bmatrix} AD_k \\ e^T \end{bmatrix}$, on obtient

$$\min_h \; c^T D_k h$$

sujet à $Bh = 0$

$$h^T h \leq \alpha^2 r^2, \quad 0 < \alpha < 1.$$

La solution analytique h^k est donnée par (voir section 4.3)

$$h^k = -\alpha r \frac{(I - B^T(BB^T)^{-1}B)D_k c}{|I - B^T(BB^T)^{-1}B\, D_k c|}.$$

D'où $y^k = e/n + h^k$ dans l'espace transformé et le nouveau point dans l'espace original $\Omega \cap S$ est donné par $x^{k+1} = T_k^{-1}(y^k)$. La vue générale de l'algorithme illustre les différentes étapes et l'utilisation des théorèmes de l'article de Karmarkar (les énoncés sont incomplets mais présentent les idées essentielles; les théorèmes sont présentés en détail à la section suivante). L'algorithme crée une séquence de point x^0, x^1, x^2,...,x^k réduisant la fonction objectif $c^T x$.

VUE GÉNÉRALE DE L'ALGORITHME

1) Problème dans l'espace original [P1] : $\min c^T x$

avec un point admissible $x^k = (x_1^k, x_2^k, \ldots x_n^k)^T > 0$

on définit $D_k = \text{diag}[x_1^k, x_2^k, \ldots, x_n^k]$.

sujet à $Ax = 0$

$e^T x = 1$

$x > 0$

2) Fonction de potentiel f(x) [P2] : $\min f(x) = \sum_{j=1}^{n} \ln \left(\frac{c^T x}{x_j}\right)$

Th.1 : à partir de $x^0 = (1/n, \ldots, 1/n)^T$, une réduction de f(x) implique une réduction de $c^T x$.

sujet à $Ax = 0$

$e^T x = 1$

$x > 0$

3) Transformation projective T_k :

$$T_k(x) = y = \frac{D_k^{-1}x}{e^T D_k^{-1}x}$$

$$T_k(x^k) = e/n$$

Th. 2 : Toute réduction de g(y) implique une réduction de f(x)

[P3] : $\min g(y) = \sum_{j=1}^{n} \ell n \left(\frac{c^T D_k y}{y_j}\right) - \sum_{j=1}^{n} \ell n\, x_j^k$

sujet à $AD_k y = 0$

$e^T y = 1$

$y > 0$

4) Approximation sur la sphère

Th. 3 : Il existe y' tel que

$$g(y') < g(e/n) - \delta_1$$

[P4] : $\min g(y)$

sujet à $AD_k y = 0$

$e^T y = 1$

$y \in B(e/n, \alpha r)$, $0 < \alpha < 1$

5) Approximation du 1er ordre sur la sphère

$y = (e/n + h) \in B(e/n, \alpha r)$, $0 < \alpha < 1$

Th. 4 : $y^k = (e/n + h^k)$ optimal est tel que

$$g(y^k) < g(e/n) - \delta_2$$

[P5] : $\min_h c^T D_k h$

sujet à $AD_k h = 0$

$e^T h = 0$

$h^T h < \alpha^2 r^2$, $0 < \alpha < 1$

Th. 5 : solution analytique pour h^k.

6) Transformation projective inverse T_k^{-1}

$$x^{k+1} = T_k^{-1}(y^k) = \frac{D_k y^k}{e^T D_k y^k} > 0, \text{ dans l'espace original.}$$

7) Critère d'arrêt

$$c^T x^{k+1} = 0 \quad \text{ou} \quad \frac{c^T x^{k+1}}{c^T x^0} < 2^{-q} .$$

4. QUELQUES DÉMONSTRATIONS

4.1 Les rayons des sphères inscrite et circonscrite

Soit $S = \{ y \in \mathbb{R}^n \mid e^T y = 1, \ y \geq 0 \}$. Le centre de ce simplexe est donné par (1/n,...,1/n) (Voir figure 1).

Le rayon R de la plus petite sphère circonscrite au simplexe, centrée en (1/n,...,1/n), est donné par la distance du centre à l'un des sommets du simplexe. Chaque sommet possède toutes ses composantes nulles, sauf une de valeur 1. D'où

$$R^2 = (n-1)\ (0-1/n)^2 + (1-1/n)^2 = \frac{n-1}{n} \text{ d'où } R = \sqrt{\frac{n-1}{n}}$$

Pour évaluer le rayon r de la plus grande sphère inscrite dans le simplexe, centrée en (1/n,...,1/n), il faudrait calculer la distance du centre au milieu d'une face (d'un côté si n=3). Un moyen plus simple est de calculer le rapport R/r. Associons à chaque sommet du simplexe, un poids de une unité. Le point (1/n,...,1/n) étant le centre de gravité du simplexe, il se situe sur le segment de droite reliant un sommet du simplexe (avec un poids de une unité) et le milieu de la face opposée (auquel on doit associer un poids de n-1 unités). Ainsi

$$1 \cdot R = (n-1) \cdot r \text{ d'où } R/r = n-1, \text{ et } r = \frac{1}{\sqrt{n(n-1)}} .$$

4.2 Approximation du 1er ordre de g(y)

$$\min_y g(y) = \min_h g(e/n + h) = \min_h g(e/n) + \nabla_y g(e/n)\cdot h \equiv \min_h \nabla_y g(e/n)\cdot h$$

$$\frac{\partial}{\partial y_j} g(y) = \frac{\partial}{\partial y_j} \left[\sum_{j=1}^{n} \ell n \left(\frac{c^T D_k y}{y_j}\right) - \sum_{j=1}^{n} \ell n\, x_j^k \right]$$

$$= \frac{\partial}{\partial y_j} [n \ell n\, c^T D_k y - \sum_{j=1}^{n} \ell n\, y_j]$$

$$= \left(\frac{n\, c_j x_j^k}{c^T D_k y} - \frac{1}{y_j}\right)$$

Évalué en y = e/n, on obtient que

$$\nabla_y g(e/n)\cdot h = \sum_{j=1}^{n} \left(\frac{\partial}{\partial y_j} g(e/n)\right) h_j$$

$$= \sum_{j=1}^{n} \left(\frac{n\, c_j\, x_j^k}{(1/n) c^T D_k e} - \frac{1}{(1/n)}\right) h_j$$

$$= \left(\frac{n}{(1/n) c^T D_k e}\right) \sum_{j=1}^{n} c_j\, x_j^k\, h_j - \frac{1}{(1/n)} \sum_{j=1}^{n} h_j$$

$$= \text{(constante)}\ c^T D_k h. \qquad \text{(puisque } e^T h = 0 \text{ dans [P5])}$$

Ainsi le point h minimisant g(e/n + h) est approché par le point h^k minimisant $c^T D_k h$ dans le problème [P5].

4.3 Théorème 5

La solution analytique du problème [P5] est donnée par

$$h^k = \alpha r \frac{v}{|v|} \quad \text{où} \quad v = -(I - B^T(BB^T)^{-1}B) D_k c .$$

Preuve : Cette démonstration, différente de celle donnée par Karmarkar est basée sur les conditions d'optimalité de Kuhn-Tucker. Une autre démonstration est aussi donnée par Minoux (1983), Programmation mathématique, Tome 1, Dunod, p. 199.

On définit, avec $h \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^{n+1}$ et $\mu \in \mathbb{R}^+$, le Lagrangien $L(h,\lambda,\mu) = c^T D_k h + \lambda^T Bh + \mu(h^T h - \alpha^2 r^2)$. À l'optimalité, $\nabla_h L = 0$, d'où

$$D_k c + B^T \lambda + 2\mu h = 0$$

$$BD_k c + BB\lambda^T + 2\mu\, Bh = 0 \quad \text{où} \quad Bh = 0$$

$$\lambda = -(BB^T)^{-1} BD_k c$$

En remplaçant dans l'équation précédente,

$$D_k c + B^T(-B(BB^T)^{-1} BD_k c) + 2\mu h = 0$$

$$h = -\frac{I - B^T(BB^T)B)D_k c}{2\mu} = \frac{v}{2\mu}$$

Pour que $|h| = \alpha r$, on pose $h^k = \alpha r \frac{v}{|v|}$.

4.4 Théorème 3

Pour le problème [P4], il existe un point y' tel que :

$c^T D_k y' = 0$ ou

$g(y') < g(e/n) - \delta_1$ où $\delta_1 = \ell n\,(1 + \frac{n\alpha}{n-1-\alpha}) > \ell n\,(1+\alpha)$.

Preuve :

Soit $y^* \in \{y \mid AD_k y = 0,\ e^T y = 1,\ y > 0\}$ le point où la fonction $c^T D_k y$ atteint son minimum, c'est-à-dire zéro ($\min c^T x = 0 \iff \min c^T D_k y = 0$).

1er cas : si $y^* \in B(e/n, \alpha r)$, alors $y' = y^*$ et $c^T D_k y' = 0$.

2e cas : sinon, on pose $y' = (1-\theta)e/n + \theta y^*$ pour une certaine valeur de $\theta \in (0,1)$ telle que y' est le point sur la frontière de la sphère $B(e/n, \alpha r)$, à l'intersection du segment de droite reliant e/n et y*. Comme combinaison convexe de e/n et y* sur la sphère, le point y' satisfait les contraintes du problème [P4]. On cherche à borner inférieurement

$$g(e/n) - g(y') = \sum_{j=1}^{n} \ell n \frac{c^T D_k e/n}{(1/n)} - \sum_{j=1}^{n} \ell n \frac{c^T D_k y'}{y'_j} = \sum_{j=1}^{n} \ell n \frac{c^T D_k e}{c^T D_k y'}\ y'_j\ .$$

Quelques résultats simples sont nécessaires à la démonstration :

1) $c^T D_k y' = (1-\Theta)\, c^T D_k e/n + \Theta\, c^T D_k y^* = (1-\Theta)\, c^T D_k e/n$

2) $y'_j = (1-\Theta)(1/n) + \Theta y^*_j$

3) si $P_i > 0$, alors $\prod_i (1+P_i) > 1 + \sum_i P_i$, d'où $\sum_i \ell n\,(1+P_i) > \ell n\,(1+\sum_i P_i)$

4) $y'-e/n = \Theta(y^*-e/n)$ d'où $\alpha r = |y'-e/n| = \Theta|y^*-e/n| < \Theta R$ et $\Theta > \frac{\alpha r}{R} = \frac{\alpha}{n-1}$

Ainsi $g(e/n)-g(b') = \sum_{j=1}^{n} \ell n\,\left(\frac{1}{(1-\Theta)} \cdot \frac{(1-\Theta)(1/n)+\Theta y^*_j}{(1/n)}\right)$ par 1) et 2)

$= \sum_{j=1}^{n} \ell n\,\left(1 + \frac{\Theta y^*_j}{(1-\Theta)(1/n)}\right) > \ell n\,\left(1 + \frac{n\Theta}{(1-\Theta)} \sum_{j=1}^{n} y^*_j\right)$ par 3)

où $\sum_{j=1}^{n} y^*_j = 1$

$= \ell n\,\left(1 + \frac{n\Theta}{(1-\Theta)}\right) > \ell n\,\left(1 + \frac{n\alpha}{n-1-\alpha}\right) > \ell n\,(1+\alpha)$.

4.5 Théorème 4

Pour le problème [P5], le point optimal y^k est tel que :

$c^T D_k y^k = 0$ ou

$g(y^k) < g(e/n) - \delta_2$ où $\delta_2 = \delta_1 - \frac{(\alpha r n)^2}{2(1-\alpha)}$

Lemme 4.1 : Pour $|x| < \alpha < 1$, alors $|\ell n(1+x)-x| < x^2/2(1-\alpha)$

Preuve :

$$\ell n(1+x) = x - \frac{x^2}{2(1+\bar{x})^2} \qquad \text{où } \bar{x} < x < \alpha$$

par le théorème de la valeur moyenne

$$|\ell n(1+x) - x| = \frac{x^2}{2(1+\bar{x})^2} < \frac{x^2}{2(1-\alpha)}$$

Lemme 4.2 : Si $y \in B(e/n, \alpha r)$ alors $1/n - \alpha/n < y_j < 1/n + \alpha/n$, $j=1,\ldots,n$.

Preuve : Sur le segment le segment de droite reliant e_j et e/n, y_j s'exprime comme $y_j = \Theta + (1-\Theta)(1/n)$, $\Theta \in [0,1]$ et la valeur maximum de y_j est atteinte sur le bord de la sphère en $\Theta = \alpha r/R = \alpha/(n-1)$. On en déduit $y_j < 1/n + \alpha/n$ et par symétrie, $y_j > 1/n - \alpha/n$.

Lemme 4.3 : Si $e^T y = 1$, $y > 0$ alors $\max \sum_{j=1}^{n} \ln y_j = -n \ln n$.

Preuve : On pose $y_1 = 1-y_2-\ldots-y_n$ pour obtenir la fonction

$\ln(1-y_2-\ldots-y_n) + \ln y_2 + \ldots + \ln y_n$, $y_j > 0$, $j=2,\ldots,n$

à maximiser. À l'optimum, la dérivée par rapport à chacune des variables est nulle, d'où :

$$\frac{-1}{1-y_2-\ldots-y_n} + \frac{1}{y_j} = \frac{-1}{y_1} + \frac{1}{y_j} = 0 \qquad j=2,\ldots,n$$

$y_1 = y_j$, $j=2,\ldots,n$ d'où $y_j = 1/n$, $j=1,2,\ldots,n$.

Ainsi, $\max \sum_{j=1}^{n} \ln y_j = n \ln (1/n) = -n \ln n$.

Lemme 4.4

Soit $y \in \{ y \in \mathbb{R}^n \mid e^T y = 1,\ y \in B(e/n, \alpha r) \}$

Alors $\dfrac{-(\alpha r n)^2}{2(1-\alpha)} < \sum_{j=1}^{n} \ln \left(\dfrac{y_j}{1/n}\right) < 0.$

Preuve :

$$y \in B(e/n, \alpha r) \Rightarrow |y-e/n|^2 < \alpha^2 r^2 \Rightarrow \sum_{j=1}^{n} \left(\frac{y_j - 1/n}{1/n}\right)^2 < \frac{\alpha^2 r^2}{(1/n)^2} = (\alpha r n)^2$$

D'autre part, $\left| \dfrac{y_j - 1/n}{1/n} \right| < \alpha$, $j=1,\ldots,n$ par le lemme 4.2.

d'où $\left| \ln \left(1 + \dfrac{y_j - 1/n}{1/n}\right) - \dfrac{y_j - 1/n}{1/n} \right| < \dfrac{1}{2(1-\alpha)} \left(\dfrac{y_j - 1/n}{1/n}\right)^2$ par le lemme 4.1.

De plus, $\sum_{j=1}^{n} \dfrac{y_j - 1/n}{1/n} = 0$

d'où $\left| \sum_{j=1}^{n} \ln \left(\dfrac{y_j}{1/n}\right) \right| = \left| \sum_{j=1}^{n} \ln \left[\left(\dfrac{y_j}{1/n}\right) - \dfrac{y_j - 1/n}{1/n}\right] \right|$

$$= \left| \sum_{j=1}^{n} \left[\ln \left(1 + \frac{y_j - 1/n}{1/n}\right) - \frac{y_j - 1/n}{1/n}\right] \right|$$

$$< \left| \sum_{j=1}^{n} \ln \left(1 + \frac{y_j - 1/n}{1/n}\right) - \frac{y_j - 1/n}{1/n} \right|$$

$$< \frac{1}{2(1-\alpha)} \sum_{j=1}^{n} \left(\frac{y_j - 1/n}{1/n}\right)^2 < \frac{(\alpha r n)^2}{2(1-\alpha)} .$$

$$\text{D'où} \quad \sum_{j=1}^{n} \ell n \left(\frac{y_j}{1/n}\right) > \frac{-(\alpha r n)^2}{2(1-\alpha)} .$$

Par le lemme 4.3 on obtient l'inégalité à zéro du côté droit.

Preuve du théorème 4.

On définit $\bar{g}(y) = \sum_{j=1}^{n} \ell n \dfrac{c^T D_k y}{c^T D_k e/n}$

et $G(y) = \bar{g}(y) + g(e/n) - g(y)$.

Soit $y' \in \{y \in \mathbb{R}^n \mid AD_k y = 0, \ e^T y = 1, \ y \in B(e/n, \alpha r)\}$ le point où $g(y)$ atteint son minimum.

$$\begin{aligned} g(e/n) - g(y^k) &= g(e/n) - g(y') + g(y') - g(y^k) \\ &= [g(e/n) - g(y')] - [\bar{g}(y') + g(e/n) - g(y')] \\ &\qquad + [\bar{g}(y^k) + g(e/n) - g(y^k)] + [\bar{g}(y') - \bar{g}(y^k)] \\ &= [g(e/n) - g(y')] - G(y') + G(y^k) + [\bar{g}(y') - \bar{g}(y^k)] \end{aligned}$$

Si le minimum de $c^T D_k y$ pour le problème [P5] est nul, alors le théorème est trivialement vrai. Sinon,

1) $g(e/n) - g(y') > \delta_1$. (par le théorème 3)

2) $$G(y) = \sum_{j=1}^{n} \ell n \frac{c^T D_k y}{c^T D_k e/n} + \sum_{j=1}^{n} \ell n \left(\frac{c^T D_k e/n}{1/n}\right) - \sum_{j=1}^{n} \ell n \left(\frac{c^T D_k y}{y_j}\right)$$

$$= \sum_{j=1}^{n} \ell n \frac{y_j}{e/n}$$

d'où, par le lemme 4.4

$$-G(y') > 0 \quad \text{et} \quad G(y^k) > \frac{-(\alpha r n)^2}{2(1-\alpha)} .$$

3) $\bar{g}(y)$ est une fonction monotone croissante de $c^T D_k y$. Le minimum sur l'espace $\{y \in \mathbb{R}^n \mid AD_k y = 0,\ e^T y = 1,\ y \in B(e/n, \alpha r)\}$ est donc atteint au même point, i.e. y^k. Ainsi,

$$\bar{g}(y') - \bar{g}(y^k) > 0.$$

De 1), 2) et 3), on obtient que

$$g(e/n) - g(y^k) > \delta_1 - \frac{(\alpha r n)^2}{2(1-\alpha)}$$

$$> \ln(1+\alpha) - \frac{(\alpha r n)^2}{2(1-\alpha)}$$

$$> \alpha - \frac{\alpha^2}{2} - \frac{\alpha^2 n}{2(n-1)(1-\alpha)}.$$

4.6 Théorème 2

Ou bien 1) $c^T x^{k+1} = 0$

ou 2) $f(x^{k+1}) < f(x^k) - \delta$.

Preuve : On obtient en appliquant la transformation projective inverse

$$T_k^{-1}(y) = \frac{D_k y}{e^T D_k y} = x$$ et les théorèmes 3 ou 4 :

1) $c^T D_k y^k = 0 \iff \dfrac{c^T D_k y^k}{e^T D_k y^k} = 0 \iff c^T x^{k+1} = 0$

2) $g(y^k) < g(e/n) - \delta \iff f(x^{k+1}) < f(x^k) - \delta$.

4.7 Théorème 1

En $k = O(nq)$ itérations, l'algorithme trouve un point admissible $x^k \in \{x \in \mathbb{R}^n \mid Ax = 0,\ e^T x = 1,\ x > 0\}$ tel que

1) $c^T x^k = 0$

ou

2) $\dfrac{c^T x^k}{c^T x^0} < 2^{-q}$, où $x^0 = (1/n, \ldots, 1/n)^T$.

Preuve. Si la condition $c^T x^k = 0$ n'est pas atteinte dans les m premières itérations, alors

$$f(x)^k \leq f(x^{k-1}) - \delta \qquad k = 1,\ldots,m$$

d'où $f(x^k) \leq f(x^0) - k\delta$

$$\sum_{j=1}^{n} \ell n \left(\frac{c^T x^k}{x_j^k}\right) \leq \sum_{j=1}^{n} \ell n \left(\frac{c^T x^0}{x_j^0}\right) - k\delta$$

$$n \ell n \left(\frac{c^T x^k}{c^T x^0}\right) \leq \sum_{j=1}^{n} \ell n \left(\frac{x_j}{1/n}\right) - k\delta \leq - k\delta \text{ (par le lemme 4.3)}$$

$$\ell n \left(\frac{c^T x^k}{c^T x^0}\right) \leq - \frac{k}{n} \delta .$$

Après $k = O(nq)$ itérations,

$$\ell n \left(\frac{c^T x^k}{c^T x^0}\right) \leq - O(q), \text{ i.e. } \frac{c^T x^k}{c^T x^0} \leq 2^{-q} .$$

5. EXTENSIONS AU MODÈLE

5.1 Discussion sur les hypothèses de départ

La présente section permet d'éliminer certaines des restrictions imposées par les hypothèses de départ de la section 1. Soit le problème standard de la programmation linéaire :

$$\min c^T x$$

$$\text{sujet à } Ax = b$$

$$x \geq 0$$

H1 : On peut toujours imposer la contrainte supplémentaire $\sum_{j=1}^{n} x_j \leq M$ où $M = 2^{O(L)}$. En ajoutant la variable d'écart x_{n+1} à l'équation précédente et en divisant par M de part et d'autre, on obtient en posant

$$\tilde{x}_j = x_j/M \quad j=1,\ldots,n+1$$
$$\tilde{c}_j = M\, c_j \quad j=1,\ldots,n; \quad \tilde{c}_{n+1} = 0$$
$$\tilde{A}_j = M\, A_j \quad j=1,\ldots,n; \quad \tilde{A}_{n+1} = 0,$$

le système suivant : $\min \tilde{c}^T\tilde{x}$

$$\text{sujet à } \tilde{A}\,\tilde{x} = b$$
$$e^T\tilde{x} = 1$$
$$\tilde{x} > 0.$$

H2 : Si le système d'équations n'est pas homogène, i.e. $b \neq 0$, alors en utilisant $e^T\tilde{x} = 1$, on a

$$\tilde{A}\,\tilde{x} = b = b(e^T\tilde{x}) \text{ d'où le système } (\tilde{A}-be^T)\,\tilde{x} = 0 \text{ qui est homogène.}$$

Les colonnes de la nouvelle matrice sont données par

$$\bar{A}_j = \tilde{A}_j - b$$
$$\bar{A}_{n+1} = -b$$

Cette matrice n'a plus la même densité que la matrice originale A si le vecteur b contient beaucoup d'éléments non nuls.

H3. Si le minimum de la fonction objectif est $\tilde{Z}_{min} \neq 0$ mais <u>connu</u>, alors on pose le nouvel objectif

$$\tilde{c}^T\tilde{x} - \tilde{Z}_{min}\, e^T\tilde{x} = (\tilde{c}^T - \tilde{Z}_{min}\, e^T)\,\tilde{x}$$

dont le minimum est zéro. Si $\tilde{Z}_{min}$ est inconnu, Karmarkar propose d'utiliser à la fois les contraintes du primal et du dual comme système d'équations (avec les modifications qui s'imposent découlant de la discussion précédente) et la valeur optimale de la <u>différence</u> de fonctions objectif primale et duale est alors nulle. Par ailleurs, Todd et Burrell [2] ont proposé une extension de la méthode de Karmarkar qui produit une séquence convergente des variables duales, permettant ainsi de traiter les problèmes où la valeur optimale de la fonction objectif est quelconque.

H4. Soit le problème suivant (obtenu par exemple par H1 et H2) :

$$\min c^T x$$

$$\text{sujet à } Ax = 0$$
$$e^T x = 1$$
$$x > 0$$

et un point admissible $x^0 > 0$, mais différent de (e/n). Par la transformation linéaire suivante, on obtient le point $\bar{x}^0 = e/n$:

$$\bar{x}_j = \frac{x_j}{n\, x_j^0} \qquad j=1,\ldots,n$$

$$\bar{c}_j = n\, x_j^0\, c_j \qquad j=1,\ldots,n$$

$$\bar{A}_j = n\, x_j^0\, A_j \qquad j=1,\ldots,n.$$

Cependant, $e^T\bar{x} = \sum_{j=1}^{n} \frac{x_j}{n\, x_j^0}$ est généralement différent de 1, et les éléments de la discussion ne permettent par d'obtenir les conditions de départ énoncées par Karmarkar. (On notera qu'une transformation linéaire ne change pas la densité de la matrice des contraintes.)

5.2 Méthode proposée

La méthode décrite ci-dessous s'inspire d'une première transformation projective proposée par Karmarkar et d'une extension de sa fonction de potentiel. D'autres méthodes sont décrites dans chacun des articles de la bibliographie.

Soit le problème standard

$$\min c^T x$$

$$\text{sujet à } Ax = b$$
$$x > 0$$

et $x^0 = (e/n)$ est un point admissible (voir dans la section précédente, la discussion H4).

1) On définit une première fonction de potentiel

$$h(x) = \sum_{j=1}^{n} \ln \frac{c^T x}{x_j} + \ln c^T x$$

et une première transformation projective donnée par

$$\bar{x}_j = \frac{x_j/x_j^0}{\sum_{j=1}^{n} (x_j/x_j^0)+1} \qquad j=1,\ldots,n$$

$$\bar{x}_{n+1} = \frac{1}{\sum_{j=1}^{n} (x_j/x_j^0)+1} .$$

La transformation inverse est donnée par

$$x_j = x_j^0 \frac{\bar{x}_j}{\bar{x}_{n+1}} \qquad j=1,\ldots,n .$$

On remarquera que $\bar{x} \in \mathbb{R}^{n+1}$, $e^T\bar{x} = 1$ et $\bar{x}^0 = e/(n+1)$. Sous cette transformation projective, $h(x)$ devient :

$$h(x) = \sum_{j=1}^{n} \ln \left(\frac{\sum_{j=1}^{n} c_j x_j^0 \bar{x}_j}{\bar{x}_j}\right) - \sum_{j=1}^{n} \ln x_j^0 + \ln \left(\frac{c_j x_j^0 \bar{x}_j}{n+1}\right)$$

$$= \bar{h}(\bar{x}).$$

Notons par $\bar{c} \in \mathbb{R}^{n+1}$ le vecteur $(c_1 x_1^0, c_2 x_1^0, \ldots, c_n x_n^0, 0)$. D'où

$$\bar{h}(\bar{x}) = \sum_{j=1}^{n+1} \ln \frac{\bar{c}^T\bar{z}}{\bar{x}_j} - \sum_{j=1}^{n} \ln x_j^0$$

$$= f(\bar{x}) - \sum_{j=1}^{n} \ln x_j^0$$

où f est la fonction de potentiel standard.

Le système d'équations Ax = b devient sous cette trasnformation :

$$\frac{A \begin{bmatrix} x_1^0 & & & \\ & x_2^0 & & \\ & & \ddots & \\ & & & x_n^0 \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_n \end{bmatrix}}{\bar{x}_{n+1}} = b \ .$$

Puisque l'algorithme de Karmarkar utilise des points admissibles $\bar{x}^0 > 0$, on peut supposer $\bar{x}_{n+1} > 0$. De plus, $x_j^0 = 1/n$, $j=1,\ldots,n$, d'où le nouveau système d'équations est

$$(1/n)\, A \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_n \end{bmatrix} - b\, \bar{x}_{n+1} = 0 \ , \text{ ou encore}$$

$\bar{A}\, \bar{x} = 0$ où $\bar{A}_j = A_j$, $j=1,\ldots,n$ et $\bar{A}_{n+1} = -nb$. Ainsi, la matrice $\bar{A}$ a une densité semblable à celle de la matrice A.

Le système à résoudre devient

$$\min f(\bar{x}) = \sum_{j=1}^{n+1} \ln \left(\frac{c^T \bar{x}}{\bar{x}_j}\right)$$

$$\text{sujet à } \bar{A}\, \bar{x} = 0$$

$$e^T \bar{x} = 1$$

$$\bar{x} > 0$$

avec un point admissible $\bar{x}^0 = e/(n+1)$.

2) On appliquera l'algorithme de Karmarkar à ce problème avec la transformation standard T en dimension n+1.

Par le théorème 4, $g(y^k) < g(e/(n+1)) - \delta$;

Par le théorème 2, $f(\bar{x}^{k+1}) < f(\bar{x}^k) - \delta$;

$\bar{h}\,(\bar{x}^{k+1}) < \bar{h}\,(\bar{x}^k) - \delta$;

$h\,(x^{k+1}) < h\,(x^k) - \delta$ en appliquant l'inverse de la première transformation projective.

3) La démonstration du théorème 1 requiert le résultat suivant :

Si $e^T x < 1$ et $x > 0$, alors $\max \sum_{j=1}^{n} \ell n\, x_j = -n\, \ell n\, n$.

Preuve : On écrit les systèmes équivalents suivants :

$$\max \sum_{j=1}^{n} \ell n\, x_j$$

$$\text{sujet à } \sum_{j=1}^{n} x_j = \Theta$$

$$0 < \Theta < 1$$

$$x_j > 0,\ j=1,\ldots,n$$

en posant $y_j = x_j/\Theta$, $j=1,\ldots,n$, on obtient le problème séparable

$$Z = \max \sum_{j=1}^{n} \ell n\, \Theta\, y = \max\ n\, \ell n\, \Theta + \sum_{j=1}^{n} \ell n\, y$$

$$\text{sujet à} \qquad 0 < \Theta < 1$$

$$\sum_{j=1}^{n} y_j = 1$$

$$y_j > 0,\ j=1,\ldots,n$$

d'où $Z = 0 - n\, \ell n\, n$ (par le lemme 4.3).

4) Preuve du théorème 1 : après k itérations, $h(x^k) < h(x^0) - k\,\delta$

$$\sum_{j=1}^{n} \ell n \frac{c^T x^k}{x_j^k} + \ell n\, c^T x^k < \sum_{j=1}^{n} \ell n \frac{c^T x^0}{x_j^0} + \ell n\, c^T x^0 - k\,\delta.$$

$$(n+1)\, \ell n \frac{c^T x^k}{c^T x^0} < \sum_{j=1}^{n} \ell n\, x_j^k - \sum_{j=1}^{n} \ell n\, x_j^0 - k\,\delta$$

en utilisant $\sum_{j=1}^{n} x_j^k < M = 2^{O(L)}$, on a

$$= \sum_{j=1}^{n} \ell n \left(\frac{x_j^k}{M} \cdot M\right) + n\, \ell n\, n - k\,\delta$$

$$= \sum_{j=1}^{n} \ell n \left(\frac{x^k}{M}\right) + n\, \ell n\, M + n\, \ell n\, n - k\,\delta$$

$$< -n\, \ell n\, n + n\, \ell n\, M + n\, \ell n\, n - k\,\delta$$

$$\text{d'où } \ell n \frac{c^T x^k}{c^T x^k} < \left(\frac{n}{n+1}\right) \ell n\, M - \left(\frac{k}{n+1}\right) \delta < \ell n\, M - \left(\frac{k}{n+1}\right) \delta$$

Avec $k = O((n+1)(L + \ell n\, L)) = O(n\, L)$, on obtient la précision désirée, i.e.

$$\frac{c^T x^k}{c^T x^0} < 2^{-O(L)} .$$

5.3 Point admissible initial

Pour appliquer l'algorithme de Karmarkar, il est nécessaire d'avoir un point admissible initial

$$x^0 \in \{x > 0 \mid Ax = b\}. \quad \text{(voir sections 5.1 (H4) et 5.2)}$$

Ce point est obtenu par une phase préliminaire (Phase I). Soit $x' > 0$ un point arbitraire. Introduisant la variable artificielle $\lambda \geq 0$, on résout le problème

$$\begin{aligned} &\min \lambda \\ &\text{sujet à} \quad Ax - (Ax' - b)\lambda = b \\ &\qquad\qquad x \geq 0, \quad \lambda \geq 0. \end{aligned}$$

Pour ce problème, le point initial $(x', \lambda = 1) > 0$ est admissible et on peut appliquer l'algorithme de Karmarkar (avec les modifications suggérées à la section précédente). Si le problème de départ est réalisable, à l'optimum $\lambda = 0$ et on obtient une solution admissible initiale x^0. Il n'est toutefois pas certain que cette solution optimale ait chacune de ses composantes strictement positives; Karmarkar suggère de résoudre le problème de la phase I jusqu'à $\lambda' < 2^{-O(L)}$. Une autre méthode d'initialisation est donnée par J.A. Tomlin [3]

6. COMPLEXITÉ

Le nombre d'itérations pour résoudre le problème est donné au théorème 1 par $O(nL)$. À chaque itération, on doit résoudre un système d'équations linéaires pour déterminer le nouveau point admissible (théorème 5), ce qui requiert $O(n^3\ L\ \ell n\ L\ \ell n\ \ell n\ L)$ opérations par la méthode d'élimination gaussienne. Toutefois, si on résout ce système d'équations en modifiant la solution de l'itération précédente, on peut économiser un facteur de $n^{\frac{1}{2}}$ (section 6 de [1]). La complexité de l'algorithme est donc $O(n^{3.5}\ L^2\ \ell n\ L\ \ell n\ \ell n\ L)$ comparée à $O(n^4\ L^2\ \ell n\ L\ \ell n\ \ell n\ L)$ pour la méthode de l'ellipsoïde.

Finalement, les résultats numériques présentés par Karmarkar et d'autres chercheurs montrent que, en pratique, le nombre d'itérations croît très lentement, soit de l'ordre de $\ell n\ n$.

7. BIBLIOGRAPHIE

[1] KARMARKAR, N., "A New Polynomial-Time Algorithm for Linear Programming". Combinatorica 4, (1984), 373-395.

[2] TODD, M.J. and BURRELL, B.P., "An Extension of Karmarkar's Algorithm for Linear Programming Using Dual Variables". School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14850, Technical Report 648 (1985).

[3] TOMLIN, J.A., "An experimental approach to Karmarkar's Projective Method for Linear Programming". Ketson Inc., Mountain View, CA 94040, (1985).

Lectures complémentaires

GHELLINCK, G. and VIAL, J.-P., "An extension of Karmarkar's algorithm for solving a system of linear homogeneous equations on the simplex", CORE paper #8538, Université Catholique de Louvain (1985).

SHANNO, D.F., "Computing Karmarkar Projections quickly", Graduate School of Administration, W.P. #85-10, University of California (1985).

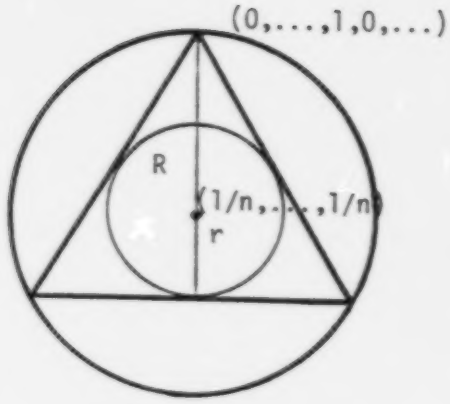

Figure 1 : Rayons des sphères circonscrites et inscrites